

520.41245X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): MORIWAKI, et al.

Serial No.: Not yet assigned

Filed: February 27, 2002

Title: PACKET COMMUNICATION APPARATUS AND CONTROLLING
METHOD THEREOF

Group: Not yet assigned



LETTER CLAIMING RIGHT OF PRIORITY

Honorable Commissioner of
Patents and Trademarks
Washington, D.C. 20231

February 27, 2002

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby
claim(s) the right of priority based on Japanese Patent Application No.(s) 2002-002318,
filed January 9, 2002.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Carl I. Brundidge
Registration No. 29,621

CIB/alb
Attachment
(703) 312-6600

日本国特許庁
JAPAN PATENT OFFICE

#4

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日

Date of Application:

2002年 1月 9日

出願番号

Application Number:

特願2002-002318

[ST.10/C]:

[JP2002-002318]

出願人

Applicant(s):

株式会社日立製作所

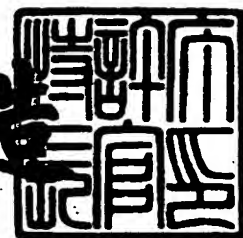


CERTIFIED COPY OF
PRIORITY DOCUMENT

2002年 2月 1日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



出証番号 出証特2002-3003808

【書類名】 特許願

【整理番号】 NT01P1079

【提出日】 平成14年 1月 9日

【あて先】 特許庁長官 殿

【国際特許分類】 H04L 12/56

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所 中央研究所内

 【氏名】 森脇 紀彦

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所 中央研究所内

 【氏名】 榊川 博史

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社日立製作所

【代理人】

 【識別番号】 100068504

 【弁理士】

 【氏名又は名称】 小川 勝男

 【電話番号】 03-3661-0071

【選任した代理人】

 【識別番号】 100086656

 【弁理士】

 【氏名又は名称】 田中 恭助

 【電話番号】 03-3661-0071

【選任した代理人】

 【識別番号】 100094352

 【弁理士】

【氏名又は名称】 佐々木 孝

【電話番号】 03-3661-0071

【手数料の表示】

【予納台帳番号】 081423

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 パケット通信装置及び該パケット通信装置を利用したパケットデータ転送制御方法

【特許請求の範囲】

【請求項 1】

複数のパケットデータを入力する入力インターフェース、該パケットデータの経路を切り替え、且つ複数の入力ポート、出力ポート及びスケジューラを備えるスイッチ部、及び切り替えられた前記経路を経由した前記パケットデータを送出する出力インターフェースから構成され、

前記入力インターフェースには、前記複数のパケットデータを格納する複数の入力バッファと、該複数の入力バッファに対応した複数の格納部と、前記入力バッファ及び格納部を制御するバッファ制御部を設け、前記入力ポートにはポートバッファを設け、

前記入力バッファの先頭に前記パケットデータが到着した場合に、前記パケットデータの内、宛先情報を含むデータ部を前記格納部に転送し、前記ポートバッファは、前記格納部から出力される前記データ部を格納し、且つ前記スケジューラにより所定の出力ポートに対し前記データ部の出力が許可されない場合、前記ポートバッファは前記データ部を廃棄し、前記バッファ制御部からの更なる出力要求に従い、前記格納部から再送される前記データ部を受け付け、前記スケジューラの指示に従って前記所定の出力ポートに対し前記データ部を転送するか、否かを決定することを特徴とするパケット通信装置。

【請求項 2】

請求項 1 に記載のパケット通信装置において、

前記バッファ制御部は、前記複数の格納部の各々に含まれる複数の前記データ部の内、1つの前記データ部を選択して、前記格納部からコピーして前記ポートバッファに出力することを特徴とするパケット通信装置。

【請求項 3】

請求項 1 に記載のパケット通信装置において、

前記スケジューラから前記所定の出力ポートに対し、前記データ部の出力許可

が為された場合、前記データ部及び後続する前記パケットデータが前記所定の出力ポートに対し切り替えられた経路を経由して転送されることを特徴とするパケット通信装置。

【請求項 4】

請求項 1 に記載のパケット通信装置であって、

前記入力インターフェースは前記入力バッファと同数のセレクタを有し、

前記セレクタは、前記バッファ制御部からの指示により、前記入力バッファの前記データ部以外の前記パケットデータ部分、或いは前記入力バッファに対応した前記格納部の前記データ部の内、何れかを選択してスイッチ部へ出力することを特徴とするパケット通信装置。

【請求項 5】

請求項 4 に記載のパケット通信装置であって、

前記入力バッファは、前記パケットデータが有する複数のセルを列形式にて格納し、前記パケットデータは先頭セル、後続セル及び最終セルを含み、前記データ部は前記先頭セル又は前記先頭セルを含む前記パケットデータであり、前記スイッチ部は前記出力ポートに対応したカウンタを備え、前記データ部は前記スイッチ部の切換えられた経路を経由し前記出力ポートにてモニタされ、前記パケットデータが有するセルの合計値が前記カウンタに格納されることを特徴とするパケット通信装置。

【請求項 6】

請求項 5 に記載のパケット通信装置であって、

前記合計値に対し前記後続セルが前記出力ポートにてモニタされる毎に前記合計値がデクリメントされ、前記カウンタ値が所定値以下に成った際に、前記出力ポートの開放を前記バッファ制御部に通知することを特徴とするパケット通信装置。

【請求項 7】

請求項 1 に記載のパケット通信装置であって、

前記入力インターフェースには、高優先度入力バッファ及び格納部並びに低優先度入力バッファ及び格納部が設けられ、前記入力バッファ及び格納部に格納さ

れる前記複数のセルを転送する際に優先度を持たせることを特徴とするパケット通信装置。

【請求項 8】

請求項 6 に記載のパケット通信装置であって、

前記セルはセルヘッダ部とセルデータ部を有し、前記後続セルの前記セルヘッダ部にはポートフリービットが設けられ、前記出力ポートは該ポートフリービットの状態に基づいて前記後続セルが通過するか、否かを判定し、前記後続セルが通過する時は、前記バッファ制御部に対し前記開放を通知することを特徴とするパケット通信装置。

【請求項 9】

請求項 8 に記載のパケット通信装置であって、

前記出力ポートは前記最終セルが通過する以前に前記バッファ制御部に対し次のパケットデータの送出を要求しうることを特徴とするパケット通信装置。

【請求項 10】

複数のパケットデータを入力する入力インターフェース、該パケットデータの経路を切り替え、且つ複数の入力ポート、出力ポート及びスケジューラを備えるスイッチ部、及び切り替えられた前記経路を経由した前記パケットデータを送出する出力インターフェースから構成され、前記入力インターフェースには複数の第 1 記憶部、該複数の第 1 記憶部に対応した複数の第 2 記憶部及び前記第 1 記憶部と第 2 記憶部を制御する制御部を設け、さらに、前記入力ポートに複数の第 3 記憶部を設けてなるパケット通信装置を利用したパケットデータ転送制御方法において、

前記第 1 記憶部に前記パケットデータが格納され、前記第 2 記憶部に前記パケットデータの宛先情報を含むデータ部が転送されるステップと、

前記第 3 記憶部は前記第 2 記憶部から出力される前記データ部を受信した後、前記スケジューラが前記データ部を選択し、所定の出力ポートに対し出力するステップとを有し、

前記スケジューラにより前記所定の出力ポートに対し前記データ部が選択されない場合には、前記第 3 記憶部が前記データ部を廃棄し、前記制御部からの更な

る出力要求に従い、前記第 2 記憶部は前記データ部を再度送出するステップとを含む事を特徴とするパケットデータ転送制御方法。

【請求項 1 1】

請求項 1 0 に記載のパケットデータ転送制御方法において、

前記パケットデータは、先頭セル、後続セル及び最終セルを含み、前記第 1 記憶部は入力キューバッファであり、前記第 2 記憶部は前記先頭セルを含む前記データ部を格納する格納バッファであり、前記第 3 記憶部はポートバッファであることを特徴とするパケットデータ転送制御方法。

【請求項 1 2】

請求項 1 0 に記載のパケットデータ転送制御方法において、

前記送出するステップにおいて、前記第 3 記憶部は前記データ部の廃棄を前記制御部に通知し、前記制御部からの前記出力要求により前記第 2 記憶部から出力された前記データ部は、前記スケジューラにより前記所定の出力ポートに転送されることを特徴とするパケットデータ転送制御方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、IP (Internet Protocol) などの可変長パケットや非同期転送モード（以下 ATM という。）の固定長パケット（一般的にセルと呼ばれる）をスイッチングするパケットデータ通信装置に関する。特に、該通信装置を利用したパケットデータ転送制御方法に関する。

【0 0 0 2】

【従来の技術】

近年、インターネットをはじめとするデータトラヒックは急激に増加している。また、従来専用線を使用して行なわれていたトランザクション処理など、高品質で、高信頼のサービスをインターネットで行おうとする動きも見られている。ネットワーク上における伝送路については波長多重技術の登場により、大容量伝送が実現されている。これに対応するため、伝送路間を接続し、パケットデータのルーティング、スイッチングを行うパケットデータ通信装置の高効率化、大容

量化，高速化が必要とされる。

【0003】

特開2000-232482号公報（以下「文献1」という。）は，クロスバスイッチを使用したパケットスイッチの一例を開示する。このパケットスイッチの構成では，入力側インタフェースに入力された可変長パケットをセル群に分割し，出力方路情報が付与された先頭セルをスイッチに入力する。スイッチはこの先頭セルの情報に従い，各出力ポートに対してのリクエストから1つを選択するようなスケジューリング処理を行って入出力ポートの接続を行った後，パケットのスイッチングを行う。

【0004】

なお入力側インタフェースには，スイッチの出力ポートの空き情報が通知されており，入力側インタフェースは，空き出力ポートに出力を希望するパケットの先頭セルをスイッチに送信する構成とする。スイッチでは入出力ポートの切り替え単位をパケット単位としている。

【0005】

つまり，ある入力インタフェースから出力インタフェースに，1つのパケットを構成するセルの出力が開始されると，そのパケットを構成する全てのセルが出力側インタフェースに到着するまでスイッチの接続を保留する。

【0006】

【発明が解決しようとする課題】

さらに，文献1では，スイッチのスケジューリング処理について，各出力ポート単位に出力を希望する入力インタフェースを1つ選択する処理を行えばよいので，従来必要とされるような複雑なスケジューリング処理は不要となり，スイッチのポート数が増加した場合や，ポート速度が高速化した場合においても，スイッチ容量の大容量化を実現しやすい。

【0007】

しかし，文献1においては，複数の入力インタフェースから，ある空き出力ポートに出力を希望する複数の先頭セルがスイッチに入力された場合には，スケジューリング処理の結果，選択されなかった先頭セルはスイッチから出力されない

。つまり、以降のスケジューリング処理で選択されるまでは、先頭セルがスイッチ内にスタックした状態となる。

【0008】

つまり、このスタック状態によって、後続する同一出力ポート宛の優先パケットや、他の空き出力ポート行きのパケットが出力を待たされてしまうHOL (Head of Line) ブロッキングと呼ばれる現象が発生して、入力されるトラヒックのパターンや負荷状況によってはスイッチの使用効率およびパケットの品質クラス制御性能が著しく低下し、スイッチにおけるスループットの低下を招き、さらに入力インターフェースからスイッチへ転送されるパケットに優先度を持たせられないという問題が生じる。

【0009】

そこで、本発明の目的は、スイッチのポート数の増加やポート速度の向上に対応しやすく、かつ、高いスループットが実現可能なパケット通信装置を提供することである。

【0010】

また、本発明の他の目的は、品質クラス制御が有効に働くパケット通信装置を提供することにある。より具体的には、スイッチのポート数の増加やポート速度の向上に対応しやすく、かつ、優先パケットが非優先パケットに対して優先して出力可能となるようなパケット通信装置を提供することである。

【0011】

【課題を解決するための手段】

前記の課題を解決するために、本発明では、入力インターフェースにおいて可変長パケットをセル群に分割し、スイッチの出力ポート別、さらには品質クラス別に用意されたキューバッファ（以下VOQと呼ぶ）に格納する。また、1つのVOQに対して1つずつ先頭セル格納レジスタを用意する。

【0012】

あるパケットがVOQの先頭に到着した場合には、そのパケットの先頭セルは、先頭セル格納レジスタに格納され、VOQの先頭には2番目のセルが待機している状態とする。入力側インターフェースは、パケットの出力方路が表示された先

頭セルを先頭セル格納レジスタから読み出してスイッチ部に入力する。スイッチでは、各入力インタフェースから入力された複数先頭セルの要求出力ポートが重なった場合には、出力ポートにつき1つの先頭セルを選択するようにスケジューリング処理を行う。

【0013】

スケジューリング処理の結果、選択された先頭セルの出力元の入力インタフェースについては、その先頭セルの後続セルがV O Qより連続的に出力されて、ひとつの packets を構成するセル全てが出力インタフェースに到着するまで、スイッチの接続を保留する。また、スケジューリングの結果、選択されなかった先頭セルの出力元の入力インタフェースについては、他の空き出力ポート宛てに出力を希望する先頭セルを、先頭セル格納レジスタからスイッチに出力して、再度スケジューリング処理に挑戦する。

【0014】

これにより、スケジューリング処理において、選択されなかった packets がスイッチにスタックすることなく、他の出力可能な packets の優先的送出が可能となるので、スループットの向上および packets の優先制御が実現可能な packets 通信装置を提供することができる。

【0015】

【発明の実施の形態】

本発明による packets 通信装置の実施例について説明する。図1は、本発明の packets 通信装置の全体構成を示す図である。この packets 通信装置は、 n 個の入出力ポートを有し、 $n \times n$ の交換を行うスイッチ10、スイッチ10と接続される入力回線インタフェース20-1～20- n 、出力回線インタフェース30-1～30- n 、および制御部60とを有する。入力回線インタフェース20-1～20- n は、入力回線40-1～40- n を収容し、入力された可変長 packets のルーティング処理や packets バッファリングを行い、固定長セルの形式でスイッチ10間へ送信処理を行なう。出力回線インタフェース30-1～30- n は、スイッチ10からセルを受信し、元の可変長 packets の形式で出力回線50-1～50- n へ出力を行う。制御部60は、制御バス60-1を通じて、ス

イッチ 1 0, 入力回線インタフェース 2 0, および出力回線インタフェース 3 0 と接続され, これらの初期設定, 障害監視などを行う。なお, 入力回線インタフェース 2 0 - 1 ~ 2 0 - n と出力回線インタフェース 3 0 - 1 ~ 3 0 - n はそれぞれ物理的に同一のカードとして実装されることが多い。

【 0 0 1 6 】

まず, 入力回線インタフェース 2 0 の構成について詳細に説明する。入力回線インタフェース 2 0 は, 入力処理部 2 1, V O Q 2 3, V O Q 制御部 2 4 を有する。図 2 を用いて, 入力処理部 2 1 の構成例を説明する。パケットデータは入力回線 4 0 を通じて装置に入力されると, 光・電気信号変換部 (O / E) 2 1 - 1 にて, 電気信号に変換される。その後, P H Y 2 1 - 2 にて S O N E T (s y n c h r o n o u s o p t i c a l n e t w o r k) フレームなどの物理レイヤ処理が行なわれる。次に, L 2 処理部 2 1 - 3 にてパケットの抽出, エラーチェックなどのレイヤ 2 処理が行なわれる。その後, 検索エンジン 2 1 - 4 にて, 宛先 I P アドレスをもとにした, 出力ポート検索, 品質クラス検索などの, レイヤ 3 処理が行なわれる。

【 0 0 1 7 】

検索処理は, 具体的には, 検索エンジン 2 1 - 4 に接続された L 3 T A B L E 2 1 - 5 を使用する。L 3 T A B L E 2 1 - 5 には, 予め宛先 I P アドレスと, 出力ポート, 品質クラス, 次の転送先の I P アドレスであるネクストホップ I P アドレスとの対応関係がテーブル形式で格納されている。検索結果はパケットのヘッダ部分に付与する。検索エンジン 2 1 - 4 では, 図 5 に示すように可変長パケットを固定長セルに分割し, それぞれのセルにセルヘッダ 2 0 0 - 1 を付与する。

【 0 0 1 8 】

図 5 の例では, 可変長パケット 1 0 0 A がセル A 1 ~ A 4 に分割される例を示している。なお, 最終セルに端数が生じた場合にはセルの空き領域に P A D が詰め込まれる。図 6 にセルフフォーマットの例を示す。セルはセルデータ 2 0 2, セルヘッダ 2 0 1 より構成される。セルヘッダ 2 0 1 は, セルの有効／無効, および先頭／中間／最終を示す C E L L 2 0 1 - 1, パケットの品質クラスを示す Q

OS 2 0 1 - 2, スイッチの宛先ポートを示すルーティング情報 RTG 2 0 1 - 3 を含む。なお, 先頭セル以外の後続セルでは, RTG 2 0 1 - 3, QOS 2 0 1 - 2 を参照せずに, 先頭セルと同一の処理をするものとし, これらの領域をデータ領域として使用しても良い。

【 0 0 1 9 】

次に図 4 を用いて, 入力回線インタフェース 2 0 (図 1) のその他の部分の機能および動作例について説明する。検索エンジン 2 1 - 4 (図 2) から出力された可変長パケット 1 0 0 A, 1 0 0 B の先頭セルからはセルヘッダ情報 2 0 0 - 1 (図 5) が取り出され, 接続線 2 5 を通じて, これを VOQ 制御部 2 4 に送信する。VOQ 制御部 2 4 はそのヘッダ情報を解析し, 可変長パケットの出力方路に対応する VOQ (2 3 A - 1 ~ 2 3 A - n のいずれか) に順次格納されるように, VOQ 2 3 のライトアドレスを指示する。VOQ 2 3 A においては, パケットはキューの先頭に来たときに, そのパケットを構成する先頭セルを先頭セル格納レジスタ 2 3 B に移動する。

【 0 0 2 0 】

つまり, キューの先頭にはパケットを構成するセカンドセル以降のセルが格納されている状態となる。各 VOQ においては, 先頭セル格納レジスタ 2 3 B に格納されている先頭セル, もしくは VOQ 2 3 A に格納されているセカンドセル以降のセルを選択して出力可能なように SEL 2 3 1 が配備される。また, 各 VOQ の読みだしは, VOQ 制御部 2 4 からの指示に従い, セレクタ 2 3 2 によって選択される。VOQ 制御部 2 4 では, スイッチの空き出力ポート情報 1 4 0 - 1 ~ 1 4 0 - n に従って, 空いている出力ポートに該当する VOQ をラウンドロビン選択などにより選択して VOQ 2 3 に読出し指示を与える。また, VOQ 制御部 2 4 には, タイマ監視部 2 4 - 3 が設けられている。タイマ監視部 2 4 - 3 では, 出力リクエスト用の先頭セルをスイッチに出力してから, そのリクエストに対するアクノリッジ (ACK) 1 3 0 - 1 が規定時間以内に返送されるかどうかを監視しており, その結果を VOQ 選択部 2 4 - 2 に通知する。

【 0 0 2 1 】

さらに, 図 4 において先頭セル A 1 - C 1 が出力リクエストとしてセレクタ 2

3 1、2 3 2 を介し、出力ポートに対し出力された時、先頭セルは先頭セル格納レジスタ 2 3 B - 1 ~ n から廃棄されず、先頭セル格納レジスタにコピーされ保持されたままと成る。

【 0 0 2 2 】

次に再び図 1 を用いて本発明のパケット通信装置のスイッチ 1 0 について詳細を説明する。スイッチ 1 0 は入力ポート毎に設けられた入力 F I F O (F i r s t I n F i r s t O u t) バッファ 1 3 - 1 ~ 1 3 - n , n × n のスイッチングを行うクロスバスイッチ 1 1 , スケジューラ 1 2 , および、出力ポート毎に設けられた出力バッファ 1 4 - 1 ~ 1 4 - n より構成される。

【 0 0 2 3 】

入力回線インタフェース 2 0 - 1 ~ n から受信したセルが、先頭セルの場合には入力 F I F O バッファ 1 3 に格納・保持される。先頭セルのセルヘッダはスケジューラ 1 2 に送られる。スケジューラ 1 2 では、入力された先頭セルの希望する出力ポートを解析して、複数の先頭セルが出力を希望する出力ポートが重なった場合には、出力ポートにつき 1 つの先頭セルを選択するようにスケジューリング処理を行う。

【 0 0 2 4 】

スケジューラ 1 2 はセルヘッダを参照して優先度の高い先頭セルを優先して転送する。また、複数のパケットが同一優先度を持つ場合には、スケジューラ 1 2 はラウンドロビン（巡回選択）によってスイッチングを行う先頭セルを選択する。スケジューリング処理の結果、出力許可を受けた先頭セルが所望の宛先に出力されるように、クロスバスイッチ 1 1 の接続を行い、先頭セルを入力 F I F O バッファ 1 3 から読み出して出力する。

【 0 0 2 5 】

また、同時に、選択された先頭セルの出力元の入力回線インタフェース 2 0 に対しては、その結果をアクノリッジ (A C K) 1 3 0 - 1 として入力回線インタフェース 2 0 に返送する。入力回線インタフェース 2 0 では、A C K 1 3 0 - 1 を受信すると V O Q 2 3 から同一パケットを構成するセカンドセル以降の残りの後続セル読み出してスイッチ 1 0 に出力する。クロスバスイッチ 1 1 では、先頭

セルの後続セルが連続的に出力されて、ひとつのパケットを構成するセル全てが出力インタフェースに到着するまで、入出力ポートの接続を保留する。

【 0 0 2 6 】

また、スケジューリングの結果、選択されなかった先頭セルの出力元の入力インタフェースにおいては、規定時間内にACK 130-1が返送されないことをVOQ制御部24内のタイマ監視部24-3（図4）で検出する。タイムオーバーの場合には、これをVOQ選択部24-2（図4）に通知し、別の空き出力ポート宛てのパケットがある場合には、このパケットの先頭セルを先頭セル格納レジスタ23Bより読み出して、スイッチ10に送信する。新たに送信された先頭セルは、入力FIFOバッファ13に格納されている先頭セルを上書きして、次のスケジューリングに再度挑戦する。スイッチの出力ポート毎に設けられた出力バッファ14-1～14-nにおいては、パケットを構成するセルが転送中であるかどうかを常時監視して、状態を空き出力ポート情報140-1～nとして、各入力回線インタフェース20-1～nのVOQ制御部24に通知する。より具体的には、セルヘッダ201に付与されているCELL201-1を監視して、有効セルの転送が終了した時点で、空き出力ポート情報140-1～nをVOQ制御部24に通知する。

【 0 0 2 7 】

最後に、図1の出力回線インタフェース30-1～nの構成について詳細に説明する。スイッチ10から出力された可変長パケット単位のセルは、出力回線インタフェース30-1～n内の出力処理部31に送信される。出力処理部31の構成を図3を用いて説明する。入力された固定長セルはL2処理部32-3にて、元の可変長パケットに組立てを行った後、レイヤ2の処理が行なわれる。

【 0 0 2 8 】

例えば、出力回線がイーサネットの場合には、次の転送先のIPアドレスであるネクストホップIPアドレスから、接続先ルータのレイヤ2アドレス（MACアドレス）を検索して付与する処理を行なう。ネクストホップIPアドレスと、接続先ルータのレイヤ2アドレスの対応は、L2TABLE32-5に格納されている。レイヤ2処理の終了後、可変長パケットはPHY32-2にて、例え

ば S O N E T フレームヘマッピング処理が行なわれ、その後電気・光信号変換部 (E/O) 3 2 - 1 にて、光信号に変換された後、出力回線 5 0 へ送出される。

【 0 0 2 9 】

次に本発明によるパケット通信装置において実際にパケットのスイッチングが行われる際のスイッチング動作例を文献 1 と比較して示す。本発明によるパケット通信装置のスイッチング例を、図 7 ～ 図 9 を用いて説明する。図 7 においては、入力回線インタフェース 2 0 - 1 に、出力回線 5 0 - 1 行きのパケット A (セル A 1, A 2, A 3, および A 4 に分割) および、出力回線 5 0 - n 行きのパケット C (セル C 1, C 2, および C 3 に分割) が入力され、入力回線インタフェース 2 0 - n に出力回線 5 0 - n 行きのパケット D (セル D 1 と D 2 に分割), 出力回線 5 0 - n 行きのパケット E (セル E 1 と E 2 に分割) が入力されている。また、出力回線 5 0 - 1 および出力回線 5 0 - n が接続されるスイッチポートはこの時点で空きポートであると仮定する。

【 0 0 3 0 】

図 8 において、出力回線インタフェース 3 0 - 1 行きのスイッチ出力ポートが空いているので、入力回線インタフェース 2 0 - 1 からは、パケット A の先頭セル A 1 がスイッチ 1 0 の入力 F I F O バッファ 1 3 - 1 に転送され、入力回線インタフェース 2 0 - n からはパケット D の先頭セル D 1 が、スイッチ 1 0 の入力 F I F O バッファ 1 3 - n にそれぞれ転送されたと仮定する。この場合、先頭セル A 1 と先頭セル D 1 は希望する出力ポートが同一であるため、スケジューラ 1 2 で選択処理が行われる。

【 0 0 3 1 】

図 9 ではスケジューラ 1 2 での選択処理の結果、先頭セル A 1 が選択され、先頭セル D 1 が選択されない場合の例を示している。入力回線インタフェース 2 0 - 1 には、後続転送が可能であることを示す A C K 1 3 0 - 1 が通知され、これに従い、セル A 2, A 3 および A 4 が連続的にスイッチに転送され所望の出力ポートに出力される。

【 0 0 3 2 】

これに対して、入力回線インタフェース 2 0 - n では、タイマ監視部 2 4 - 3

(図4)によりACK130-nが一定時間内に通知されないことを検出すると、リクエストが認められなかったと判断し、他の空き出力ポートに対して、パケットEの先頭セルE1を、入力FIFOバッファ13-nに転送して、再度スケジューラ12での選択処理を行う。

【0033】

次に文献1におけるパケットのスイッチング例について、図10を用いて説明する。図10において、出力回線インタフェース30-1行きのスイッチ10における出力ポートが空いていると仮定した場合、入力回線インタフェース20-1からは、パケットAの先頭セルA1がスイッチ10の入力FIFOバッファ13-1に転送され、入力回線インタフェース20-nからはパケットDの先頭セルD1が、スイッチ10の入力FIFOバッファ13-nにそれぞれ転送されるとする。

【0034】

図10では、スケジューラ12での選択処理の結果、先頭セルA1が選択されるものとする、先頭セルD1は長時間スイッチにスタックする可能性がある。

【0035】

さらに、先述した通り後続する同一出力ポート宛の優先パケットや、他の空き出力ポート行きのパケットが出力を待たされてしまうHOL(Head of Line)ブロッキングと呼ばれる現象が発生する。その結果、入力されるトラヒックのパターンや負荷状況によってはスイッチの使用効率およびパケットの品質クラス制御性能が著しく低下する場合がある。

【0036】

対照的に前述した図7～図9に示す本発明によるパケット通信装置においては、スケジューリング処理にて選択されなかったパケットが入力FIFOバッファから廃棄されるので、スイッチにスタックするという問題が解決される。さらに、他の出力可能なパケットの優先的送出が可能となるので、スループットの高いパケット通信装置を提供することができる。

【0037】

本発明の第2の実施例として、スイッチ部において品質クラス制御を行う形態

について図 1 1 を用いて説明する。図 1 1 に示す品質制御対応回線インタフェースにおいては、図 4 に示した回線インタフェース 2 0 との差分のみを示す。品質制御対応回線インタフェースにおいては、V O Q 2 3 内に n 本の出力方路対応かつ 2 クラスの品質クラス別の V O Q（高優先 2 3 A H - 1 ~ 2 3 A H - n，低優先 2 3 A L - 1 ~ 2 3 A L - n）を有する。入力処理部 2 1 より転送されたパケットは、図 6 に示すセルヘッダ 2 0 1 内の、R T G 2 0 1 - 3 および Q O S 2 0 1 - 2 に従って該当する V O Q に入力される。V O Q 制御部 2 4 は、例えば、高優先キューを優先したラウンドロビン選択を行う。具体的には、ある方路への読出し指示を受け取ると、S E L 2 3 2 により方路選択を行うと同時に品質クラスセクタ S E L 2 3 1 により、高優先 V O Q 2 3 A H - x（x は 1 から n のいずれかを示す）にパケットが存在する場合にはそれを読み出し、存在しない場合には、低優先 V O Q 2 3 A L - x よりパケットの読み出しを行う。

【 0 0 3 8 】

本実施例に従うと、ある出力ポート宛ての、例えば 2 3 A L - 1 ~ 2 3 A L - n から成る低優先パケットがスイッチ 1 0（図 1）の入力 F I F O バッファ 1 3 に格納され、スケジューラ 1 2 で選択されなかった場合に、後に到着した 2 3 A H - 1 ~ 2 3 A H - n から成る高優先パケットを再送可能になる。つまり、高優先パケットが低優先パケットによりブロッキングされることが防止できるので、動画像通信、トランザクション処理などに必要とされる高品質サービスに対応可能なパケット通信装置が提供できる。

【 0 0 3 9 】

次に、本発明の第 3 の実施例について説明する。前述した第 1 及び第 2 の実施例では、出力リクエスト時に、先頭セル格納レジスタ 2 3 B - 1 ~ n に格納されている先頭セルのみをスイッチ 1 0 の入力 F I F O バッファ 1 3 に転送する例、および先頭セルのみを再送する例を示した。

【 0 0 4 0 】

しかし、実装条件によっては A C K 1 3 0 が V O Q 制御部 2 4 に返送されて、後続パケットが出力されるまで、タイムラグが生じる場合がある。そこで、このタイムラグを解消するため、先頭セルのみならず先頭セルから m セル（以下、先

頭セル群と呼ぶ)を単位として、上記と同様の処理を行う例について以下に説明する(ただし、 $m \geq 2$)。

【0041】

具体的には、図12に示すように、入力回線インタフェース20において、VOQ23A-1~nのそれぞれに対応した先頭セル群レジスタ23C-1~nを用意しておき、パケットがVOQ23A-1~nの先頭に到着すると、そのパケットの先頭セル群を先頭セル群格納レジスタ23C-1~nに転送する。先頭セル群格納レジスタ23C-1~nはmセル分を格納するレジスタである。

【0042】

図12は、一例として $m=2$ の場合を示している。出力リクエスト時には、各入力回線インタフェース20において出力可能なパケットの先頭セル群を1つ選択して、これらを全てスイッチ10(図1)の入力FIFOバッファ13へ送信する。

【0043】

スケジューリング処理の結果、出力許可を受けた先頭セル群を読み出して出力し、それと同時に、後続セルを連続的に出力するように、入力回線インタフェース20に指示する。また、出力許可を受けなかった入力インタフェース20において、他の空き出力ポート宛てに出力を希望するパケットがある場合には、そのパケットの先頭セル群をスイッチ10の入力FIFOバッファ13に出力して、以前の先頭セル群を上書きし、再度スケジューリング処理を行う。

【0044】

なお、mの値は、スケジューリングの結果、出力が許可された時点から、VOQ23Aに格納されている後続セルが入力FIFOバッファ13に到達するまでの時間がmセル時間となるように決めればよい。これにより、入力回線インタフェース20が出力許可を受けた場合、後続パケットが入力FIFOバッファ13に格納されている先頭セル群に追いつくことができる。

【0045】

よって、タイムラグのない連続したセルの形式でパケットのスイッチングが行われるため、スイッチのスループットが向上する。さらに本実施例を拡張する形

として、出力リクエスト時および再送時の単位を、先頭セル群単位ではなく 1 パケット単位とする方法も考えられる。

【 0 0 4 6 】

最後に、本発明の第 4 の実施例について説明する。前記の実施例では、スイッチの出力ポート毎に設けられた図 1 に示す出力バッファ 1 4 - 1 ~ n において、セルが転送中であるかどうかを常時監視して、空き出力ポート情報 1 4 0 - 1 ~ n を V O Q 制御部 2 4 に通知している。

【 0 0 4 7 】

しかし、実装条件によっては、空き出力ポート情報 1 4 0 - 1 ~ n が V O Q 制御部 2 4 に通知されてから、先頭セルが出力され、次の出力リクエストがスケジューラにエントリされるまでに時間がかかるため、出力回線に無駄な空きタイミング（例えば j セル時間）が生じる場合がある。

【 0 0 4 8 】

そこで、この空きタイミングをなくすため、パケットを構成する最終セルがクロスバスイッチ 1 1 を通過し終わる以前に、出力ポートを空きと判定する構成とする。具体的な方法としては、図 1 3 に示すように、セルヘッダ 2 0 1（図 6）のセルフフォーマットにポートフリービット（P F）2 0 1 - 4 を定義する（“1”：ポート使用中，“0”：ポート開放）。入力処理部 2 1（図 1、図 1 2）では、パケットをセルに分割するときに、パケットの最終セルを含む最後の j セルについては、P F 2 0 1 - 4 = “0” に設定し、それ以外のセルを P F 2 0 1 - 4 = “1” に設定する。出力バッファ 1 4 - 1 ~ n においては、ポート開放ビット P F 2 0 1 - 4 に従って、P F 2 0 1 - 4 = “0” を検出した場合に、ポートの開放を V O Q 制御部 2 4 に通知すれば良い。

【 0 0 4 9 】

上述した如く空きタイミングを無くす為に、ポートフリービットがセルヘッダに設定される。このポートフリービットの設定により、従来のパケットデータ転送に比べ、本実施例によるパケットデータ転送が提供する効果を以下に説明する。

【 0 0 5 0 】

まず従来におけるパケットデータ転送例を説明する。図 1 6 (a) に示す様に先頭パケット 1 6 0 1 は、入力回線インターフェース 2 0 (図 1) から出力されスイッチ部 1 0 に到達していた。ここで先頭パケット 1 6 0 1 は先頭セル 1 6 0 2 及び最終セル 1 6 0 3 を含んだ後続セル 1 6 0 5 から構成される。当該先頭パケット 1 6 0 1 がスイッチ部 1 0 の出力ポート 1 4 - n を通過する際、最終セル 1 6 0 3 をモニタしてその出力ポート 1 4 - n がフリーに成ったことを V O Q 制御部 2 4 (図 1) に通知していた。V O Q 制御部 2 4 はその通知信号に応じて次のパケット 1 6 0 4 を送出していたが、当該先頭パケット 1 6 0 1 の送出から次のパケット 1 6 0 4 の送出まではタイムラグが生じるという問題が発生した。

【 0 0 5 1 】

次に、本実施例によるパケットデータ転送例を説明する。図 1 6 (b) に示す様に先頭パケット 1 6 0 6 の内、最終セル 1 6 0 9 以前の後続セル 1 6 0 8 にポートフリービットを付加する。ここで当該先頭パケット 1 6 0 6 は先頭セル 1 6 0 7 及び最終セル 1 6 0 9 を含んだ後続セル 1 6 0 8 から構成される。

【 0 0 5 2 】

これにより、出力ポート 1 4 - n におけるポートフリービット付セル 1 6 0 8 をモニタできる。従って、ポートフリービットを参照することにより、実際の最終パケットが通過する以前に、V O Q 制御部 2 4 に対して次のパケットの送出を要求することが可能となる。

【 0 0 5 3 】

検出信号は V O Q 制御部 2 4 に通知される。V O Q 制御部 2 4 の指示に基づく次のパケット 1 6 1 0 の送出が迅速化し、当該先頭パケット 1 6 0 6 と次のパケット 1 6 1 0 との間のタイムラグが図 1 6 (a) に示す従来例の場合より短縮されるという効果が有る。ここで、図 1 6 (a) , (b) に示す先頭パケット 1 6 0 1 、 1 6 0 6 及び次パケット 1 6 0 4 、 1 6 1 0 において、F I R S T 及び L A S T は先頭セルと最終セルを示している。

【 0 0 5 4 】

また、別の方法としては、図 1 4 に示すように、セルヘッダ 2 0 1 のセルフフォーマットにトータルセル数情報ビット (T C N) 2 0 1 - 5 を定義する。入力処

理部 2 1 (図 1) では、パケットをセルに分割するとき、パケット長から分割セル数を計算して、セル数を先頭セルの TCN 2 0 1 - 5 にライトする。

【 0 0 5 5 】

スイッチ部 1 0 においては、図 1 5 に示すように、出力バッファ 1 4 - 1 ~ n に接続されたセル数カウンタ 1 5 - 1 ~ n を設けておく。出力バッファ 1 4 - 1 ~ n においては、先頭セルを検出すると、TCN 2 0 1 - 5 ビットから分割セル数を取得してこれをセル数カウンタ 1 5 - 1 ~ n に格納する。そして、後続セルの到着毎に、セル数カウンタ 1 5 - 1 ~ n をデクリメントし、カウンタ値が j 以下になった時点で、ポートの開放を VOQ 制御部 2 4 に通知し、セル数カウンタ 1 5 - 1 ~ n をリセットする構成とする。なお、TCN 2 0 1 - 5 は先頭セルについてのみ付与すれば良い。

【 0 0 5 6 】

以上に本発明の第 1 から第 4 の実施例に基づくパケット通信装置の特徴について述べたが、さらに上述した第 1 から第 4 の実施例に示すパケット通信装置は、以下に示す項目 (a) ~ (i) の特徴点を有するパケット通信装置として提供することも可能である。

【 0 0 5 7 】

(a) 複数のパケットデータを入力する入力インターフェース、該パケットデータの経路を切り替え、且つ複数の入力ポート、出力ポート及びスケジューラを備えるスイッチ部、及び切り替えられた前記経路を経由した前記パケットデータを送出する出力インターフェースから構成され、前記入力インターフェースには、前記複数のパケットデータを格納する複数の入力バッファと、該複数の入力バッファに対応した複数の格納部と、前記入力バッファ及び格納部を制御するバッファ制御部を設け、前記入力ポートにはポートバッファを設け、前記入力バッファの先頭に前記パケットデータが到着した場合に、前記パケットデータの内、宛先情報を含むデータ部を前記格納部に転送し、前記ポートバッファは、前記格納部から出力される前記データ部を格納し、且つ前記スケジューラにより所定の出力ポートに対し前記データ部の出力が許可されない場合、前記ポートバッファは前記データ部を廃棄し、前記バッファ制御部からの更なる出力要求に従い、前記

格納部から再送される前記データ部を受け付け、前記スケジューラの指示に従って前記所定の出力ポートに対し前記データ部を転送するか、否かを決定することを特徴とするパケット通信装置。

【 0 0 5 8 】

(b) 前記バッファ制御部は、前記複数の格納部の各々に含まれる複数の前記データ部の内、1つの前記データ部を選択して、前記格納部からコピーして前記ポートバッファに出力することを特徴とする、上記(a)に記載のパケット通信装置。

【 0 0 5 9 】

(c) 前記スケジューラから前記所定の出力ポートに対し、前記データ部の出力許可が為された場合、前記データ部及び後続する前記パケットデータが前記所定の出力ポートに対し切り替えられた経路を経由して転送されることを特徴とする、上記(a)に記載のパケット通信装置。

【 0 0 6 0 】

(d) 前記入力インターフェースは前記入力バッファと同数のセレクタを有し、前記セレクタは、前記バッファ制御部からの指示により、前記入力バッファの前記データ部以外の前記パケットデータ部分、或いは前記入力バッファに対応した前記格納部の前記データ部の内、何れかを選択してスイッチ部へ出力することを特徴とする、上記(a)に記載のパケット通信装置。

【 0 0 6 1 】

(e) 前記入力バッファは、前記パケットデータが有する複数のセルを列形式にて格納し、前記パケットデータは先頭セル、後続セル及び最終セルを含み、前記データ部は前記先頭セル又は前記先頭セルを含む前記パケットデータであり、前記スイッチ部は前記出力ポートに対応したカウンタを備え、前記データ部は前記スイッチ部の切換えられた経路を経由し前記出力ポートにてモニタされ、前記パケットデータが有するセルの合計値が前記カウンタに格納されることを特徴とする、上記(d)に記載のパケット通信装置。

【 0 0 6 2 】

(f) 前記合計値に対し前記後続セルが前記出力ポートにてモニタされる毎に

前記合計値がデクリメントされ、前記カウンタ値が所定値以下に成った際に、前記出力ポートの開放を前記バッファ制御部に通知することを特徴とする、上記（e）に記載のパケット通信装置。

【 0 0 6 3 】

（g）前記入力インターフェースには、高優先度入力バッファ及び格納部並びに低優先度入力バッファ及び格納部が設けられ、前記入力バッファ及び格納部に格納される前記複数のセルを転送する際に優先度を持たせることを特徴とする、上記（a）に記載のパケット通信装置。

【 0 0 6 4 】

（h）前記セルはセルヘッダ部とセルデータ部を有し、前記後続セルの前記セルヘッダ部にはポートフリービットが設けられ、前記出力ポートは該ポートフリービットの状態に基づいて前記後続セルが通過するか、否かを判定し、前記後続セルが通過する時は、前記バッファ制御部に対し前記開放を通知することを特徴とする、上記（f）に記載のパケット通信装置。

【 0 0 6 5 】

（i）前記出力ポートは前記最終セルが通過する以前に前記バッファ制御部に対し次のパケットデータの送出を要求しうることを特徴とする、上記（h）に記載のパケット通信装置。

【 0 0 6 6 】

しかも、上述したパケット通信装置を利用してパケットデータの転送を制御する方法として、以下の項目（i）～（i i i）の特徴を有するパケットデータ転送制御方法が提供可能である。

【 0 0 6 7 】

（i）複数のパケットデータを入力する入力インターフェース、該パケットデータの経路を切り替え、且つ複数の入力ポート、出力ポート及びスケジューラを備えるスイッチ部、及び切り替えられた前記経路を経由した前記パケットデータを送出する出力インターフェースから構成され、前記入力インターフェースには複数の第1記憶部、該複数の第1記憶部に対応した複数の第2記憶部及び前記第1記憶部と第2記憶部を制御する制御部を設け、さらに、前記入力ポートに複数

の第 3 記憶部を設けてなるパケット通信装置を利用したパケットデータ転送制御方法において、前記第 1 記憶部に前記パケットデータが格納され、前記第 2 記憶部に前記パケットデータの宛先情報を含むデータ部が転送されるステップと、前記第 3 記憶部は前記第 2 記憶部から出力される前記データ部を受信した後、前記スケジューラが前記データ部を選択し、所定の出力ポートに対し出力するステップとを有し、前記スケジューラにより前記所定の出力ポートに対し前記データ部が選択されない場合には、前記第 3 記憶部が前記データ部を廃棄し、前記制御部からの更なる出力要求に従い、前記第 2 記憶部は前記データ部を再度送出するステップとを含む事を特徴とするパケットデータ転送制御方法。

【 0 0 6 8 】

(i i) 前記パケットデータは、先頭セル、後続セル及び最終セルを含み、前記第 1 記憶部は入力キューバッファであり、前記第 2 記憶部は前記先頭セルを含む前記データ部を格納する格納バッファであり、前記第 3 記憶部はポートバッファであることを特徴とする、上記 (i) に記載のパケットデータ転送制御方法。

【 0 0 6 9 】

(i i i) 前記送出するステップにおいて、前記第 3 記憶部は前記データ部の廃棄を前記制御部に通知し、前記制御部からの前記出力要求により前記第 2 記憶部から出力された前記データ部は、前記スケジューラにより前記所定の出力ポートに転送されることを特徴とする、上記 (i) に記載のパケットデータ転送制御方法。

【 0 0 7 0 】

【発明の効果】

以上説明した本発明の実施例によれば、次のような効果が期待できる。

(1) 大容量のパケット通信装置を構成する場合に、少ないハード構成で、スケジューリング処理におけるネックの少ないスイッチを提供することができる。より具体的には、スイッチのポート数の増加やポート速度の向上に対応しやすく、かつ、高いスループットが実現可能なパケット通信装置が提供できる。

(2) 品質クラス制御が有効に働く大容量のパケット通信装置を提供することができる。

【図面の簡単な説明】

【図 1】

本発明の packets 通信装置の全体構成を示すブロック図である。

【図 2】

本発明の packets 通信装置の入力回線インタフェースカードの構成を示すブロック図である。

【図 3】

本発明の packets 通信装置の出力回線インタフェースカードの構成を示すブロック図である。

【図 4】

本発明の packets 通信装置の回線インタフェースカードの構成を示すブロック図である。

【図 5】

本発明の packets 通信装置内部でのセル分割の動作を示すブロック図である。

【図 6】

本発明の packets 通信装置で使用する、セルフフォーマット図である。

【図 7】

本発明の packets 通信装置のスイッチング動作を示す説明図である。

【図 8】

本発明の packets 通信装置のスイッチング動作を示す説明図である。

【図 9】

本発明の packets 通信装置のスイッチング動作を示す説明図である。

【図 1 0】

従来の大容量 packets スイッチのスイッチング動作を示す説明図である。

【図 1 1】

本発明の packets 通信装置の回線インタフェースカードにおける他の構成を示すブロック図である。

【図 1 2】

本発明の packets 通信装置の回線インタフェースカードにおける他の構成を示

すブロック図である。

【図 1 3】

本発明のパケット通信装置で使用する、他のセルフフォーマット図である。

【図 1 4】

本発明のパケット通信装置で使用する、他のセルフフォーマット図である。

【図 1 5】

本発明のパケット通信装置のスイッチにおける他の構成を示すブロック図である。

【図 1 6】

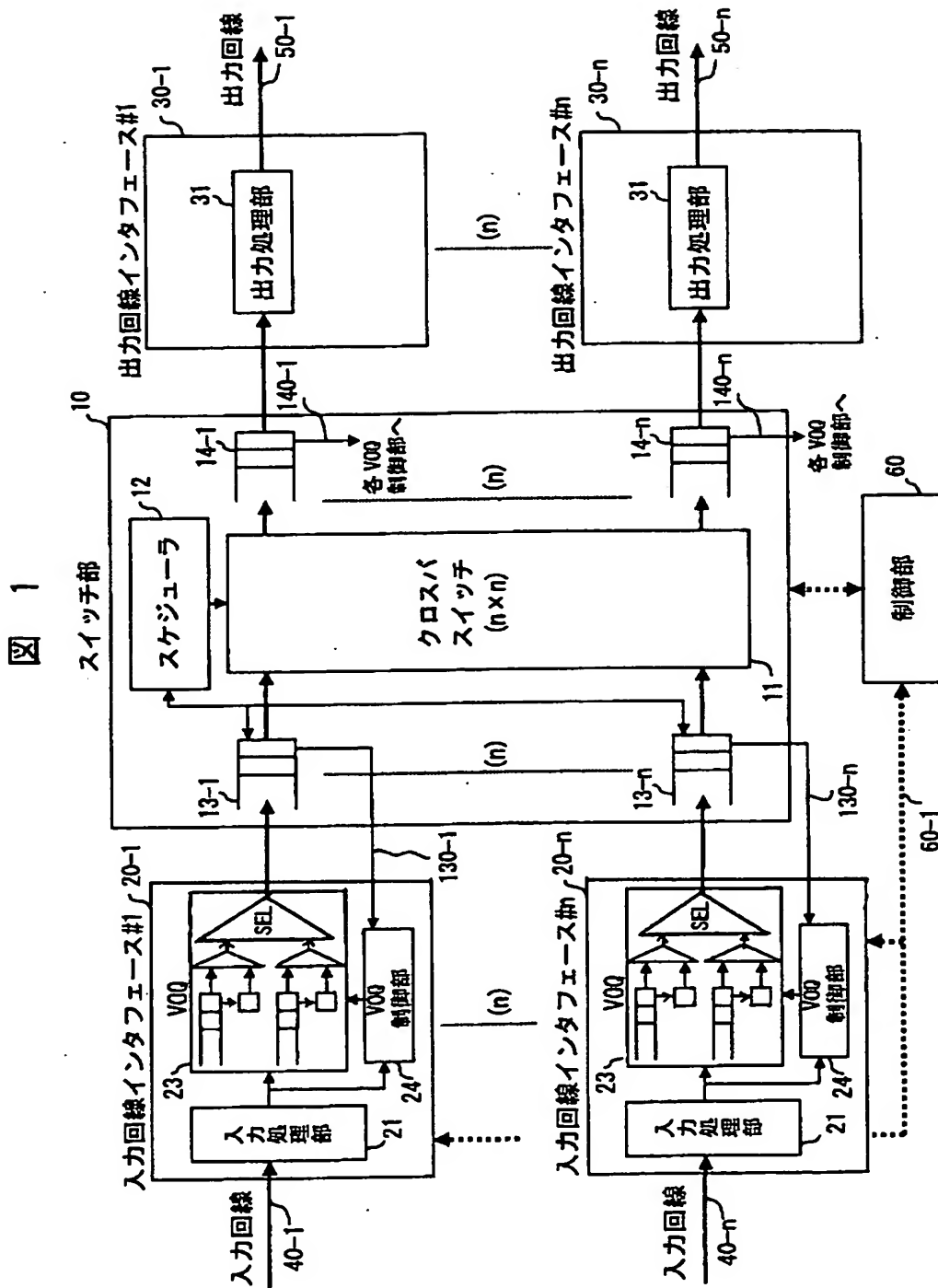
図 1 6 (a) は従来におけるパケットデータ転送の際、先頭パケットと次のパケットとの間で生じるタイムラグを示し、図 1 6 (b) は本発明のポートフリービットを付加した場合の先頭パケットと次のパケットとの間で生じるタイムラグを示す図である。

【符号の説明】

1 1 …クロスバススイッチ, 1 2 …スケジューラ, 2 0 …入力回線インタフェース, 3 0 …出力回線インタフェース, 2 3 …VOQ, 2 4 …VOQ制御部, 6 0 …制御部、先頭セル格納レジスタ…2 3 B-1 ~ n、入力FIFOバッファ…1 3 -1 ~ n、セクタ…2 3 1 -1 ~ n。

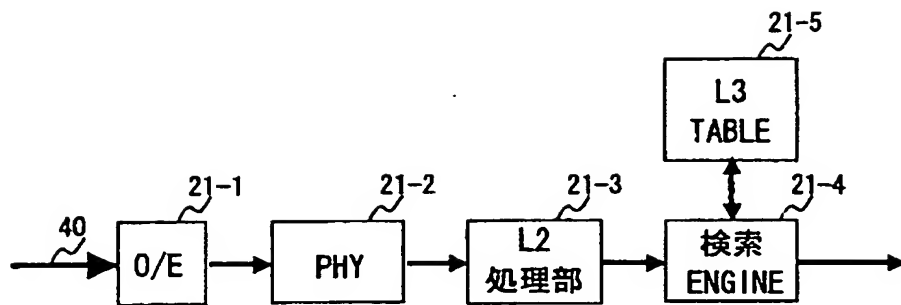
【書類名】 図面

【図 1】



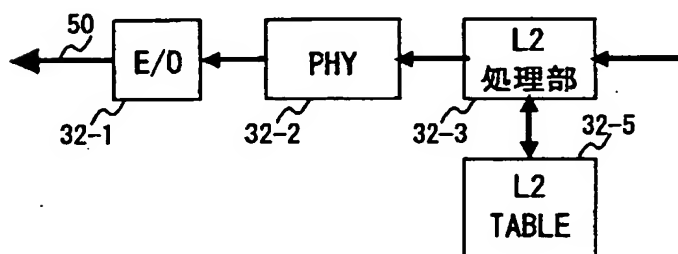
【图 2】

图 2

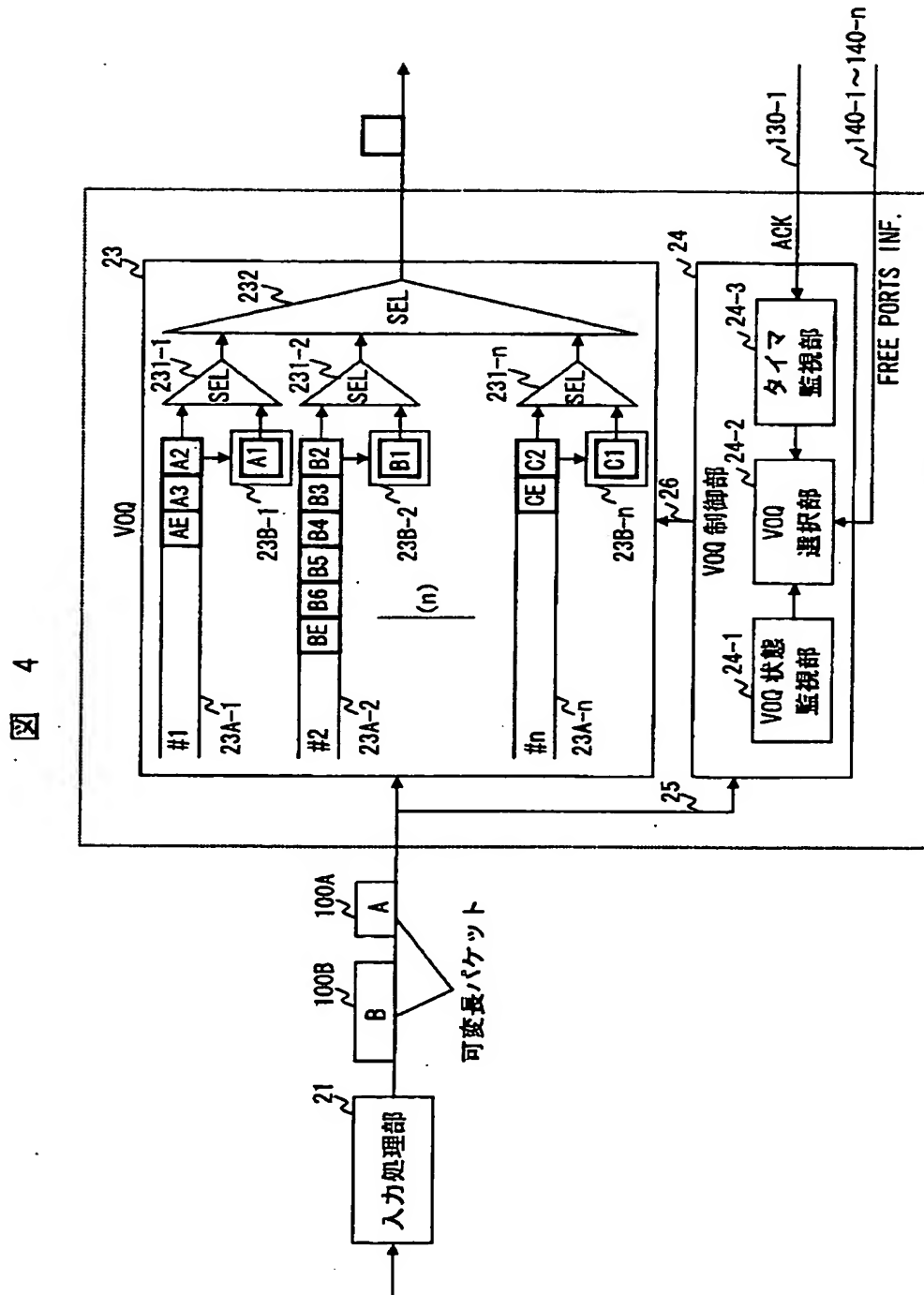


【图 3】

图 3

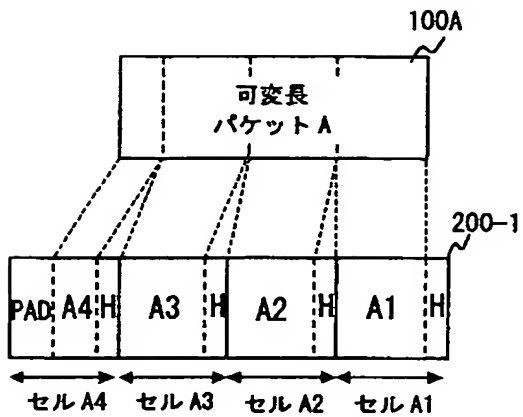


【図 4】



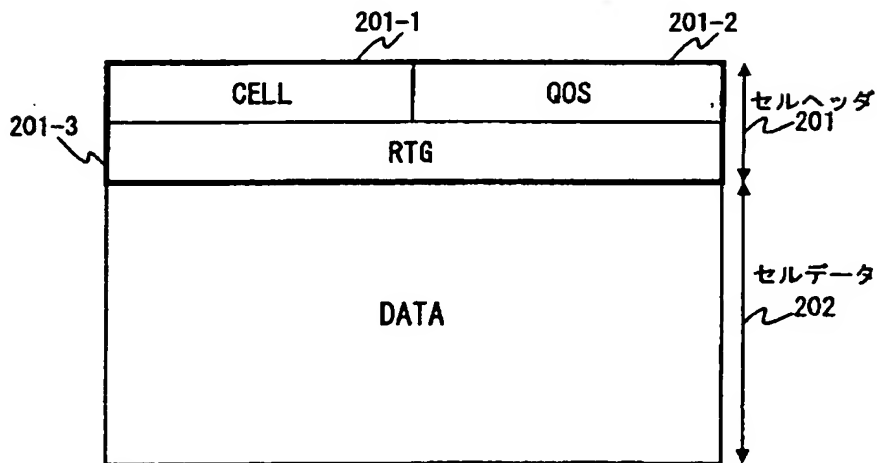
【図 5】

図 5



【図 6】

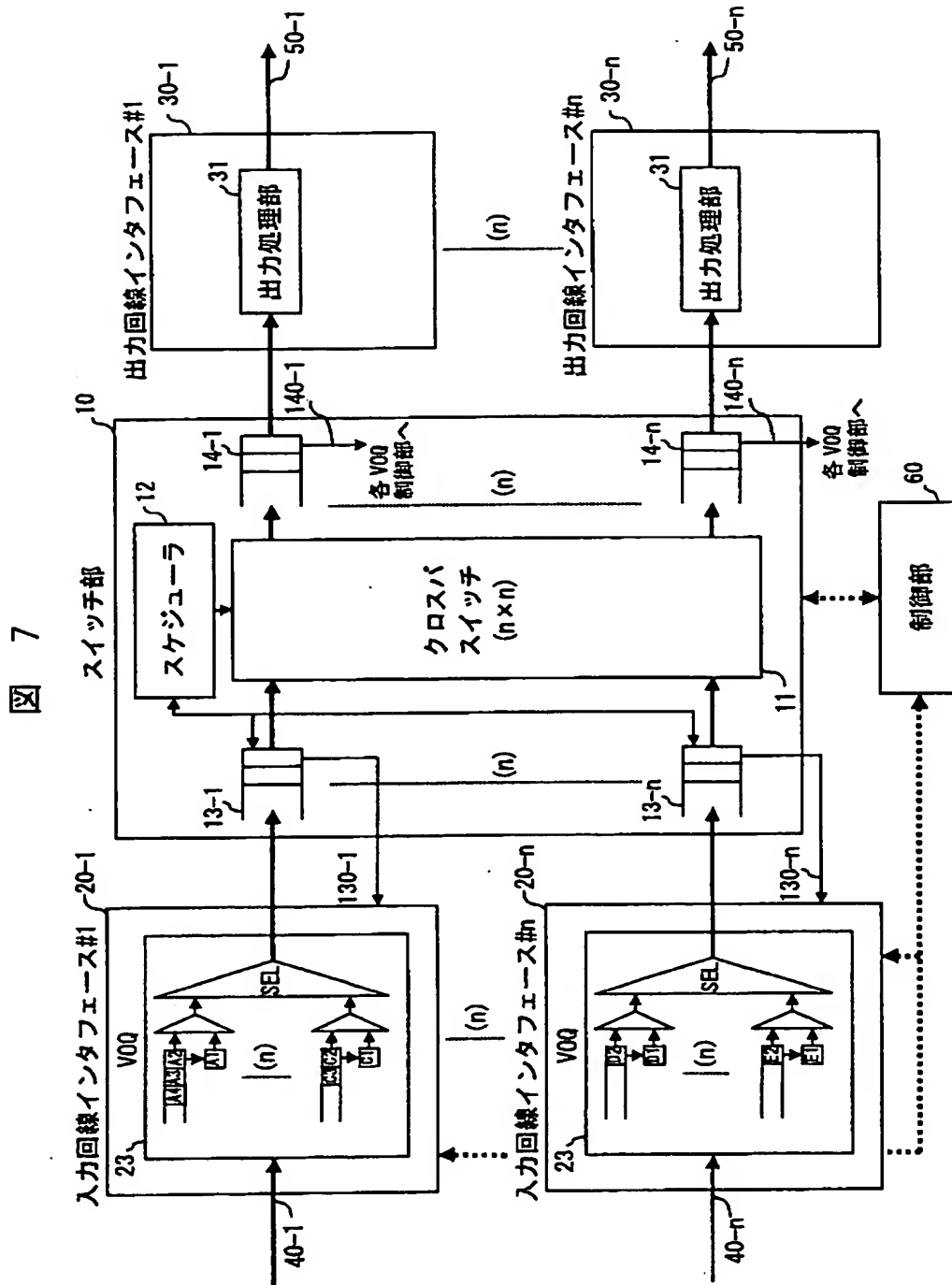
図 6



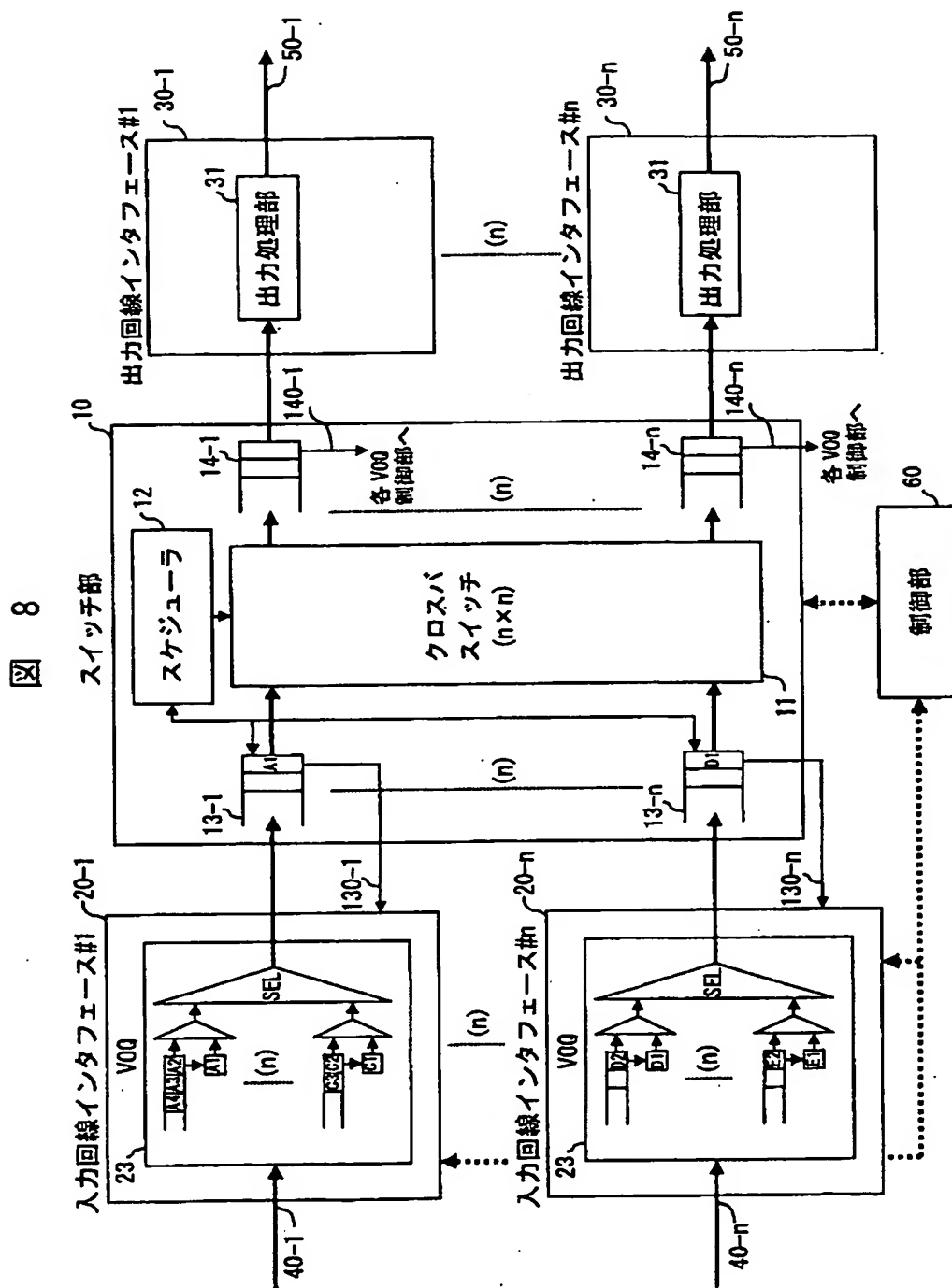
CELL bit

- 0xx : 無効セル
- 110 : 先頭セル
- 100 : 中間セル
- 101 : 最終セル

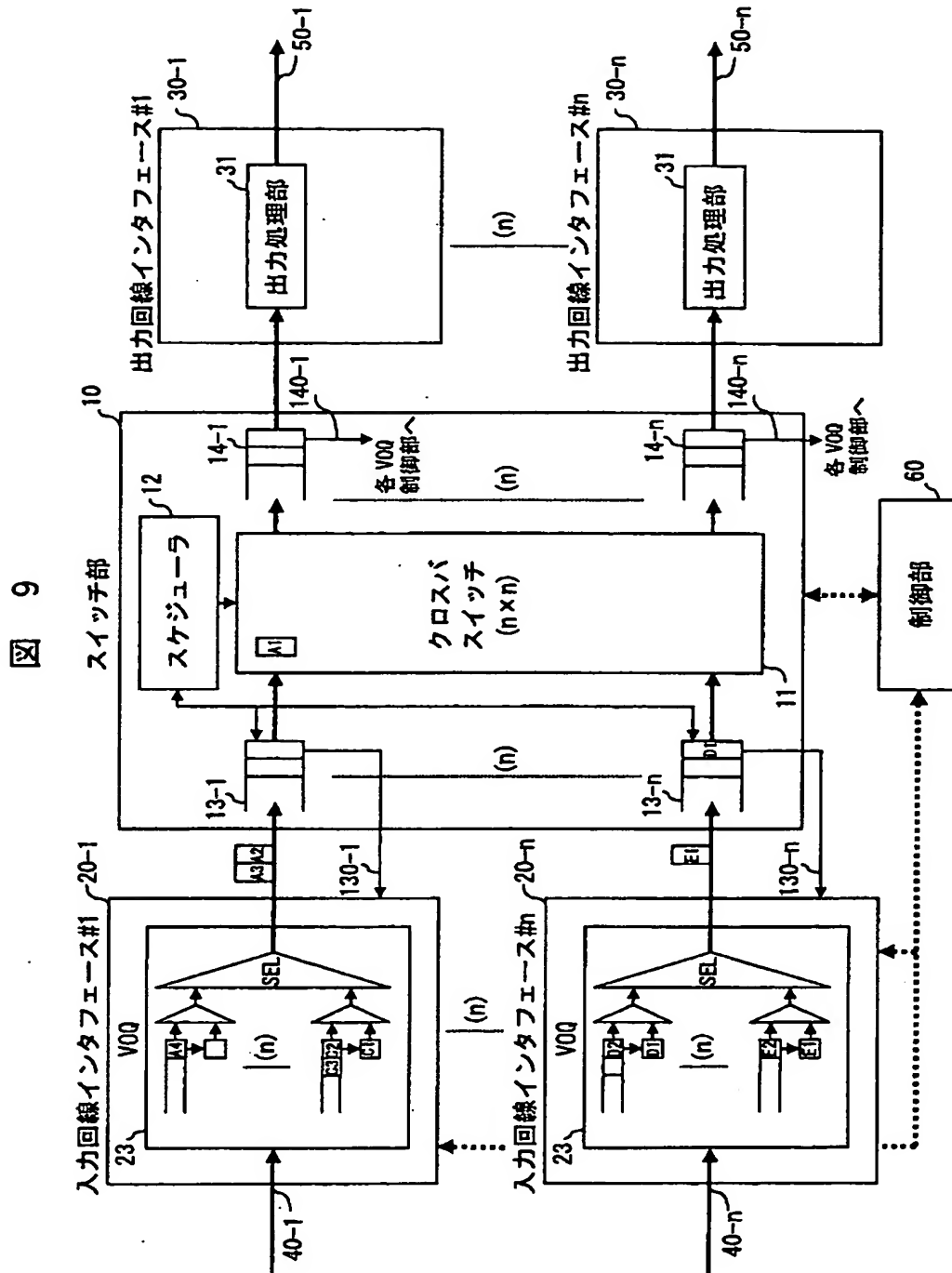
【図 7】



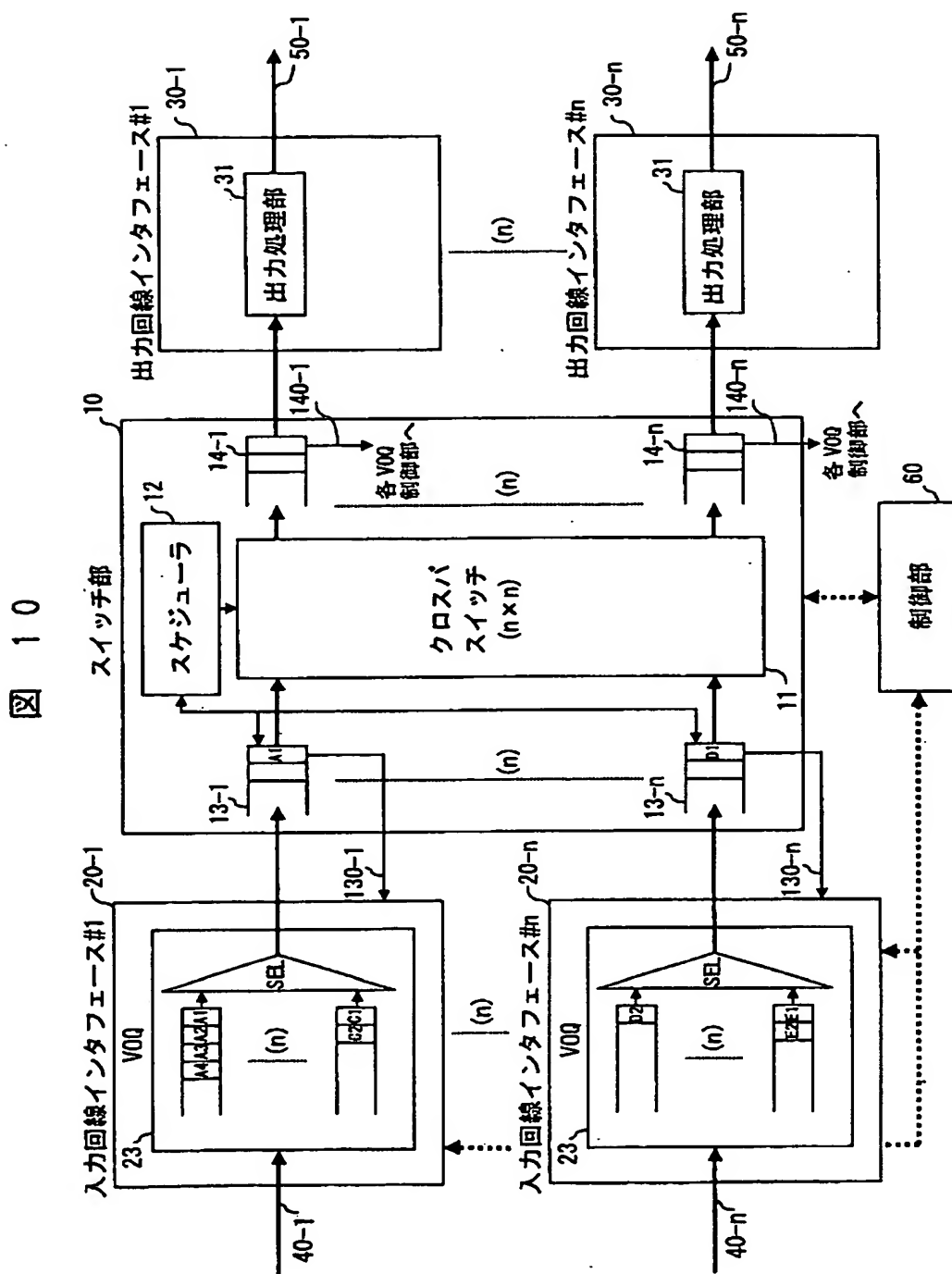
【図8】



【図9】

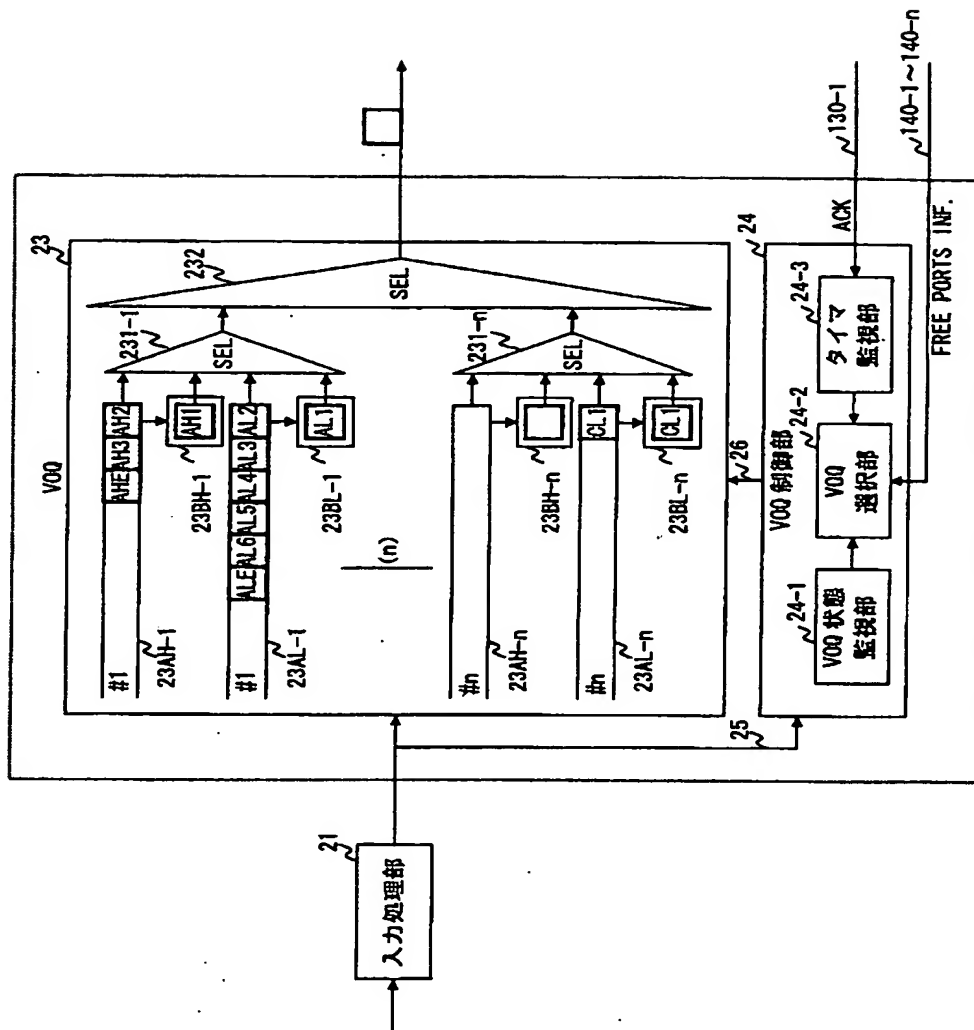


【図10】

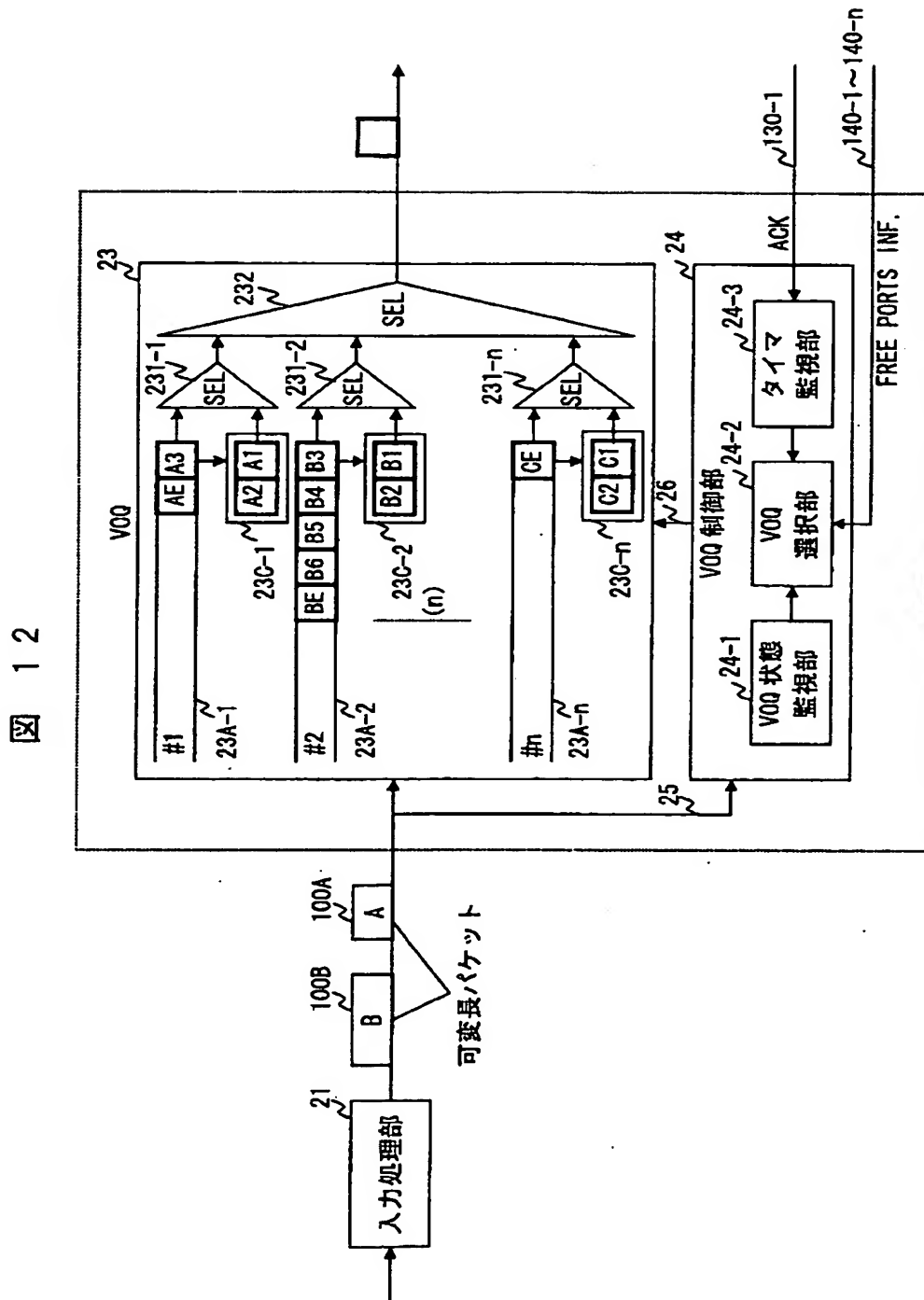


【図 11】

図 11

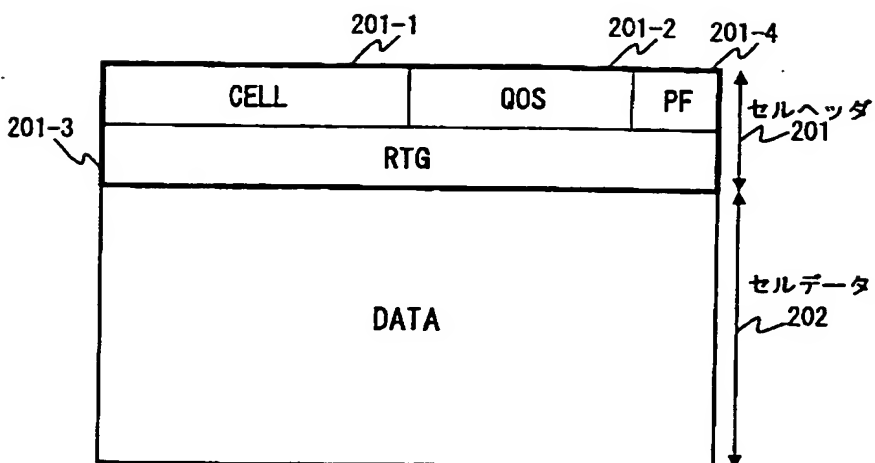


【图 1 2】



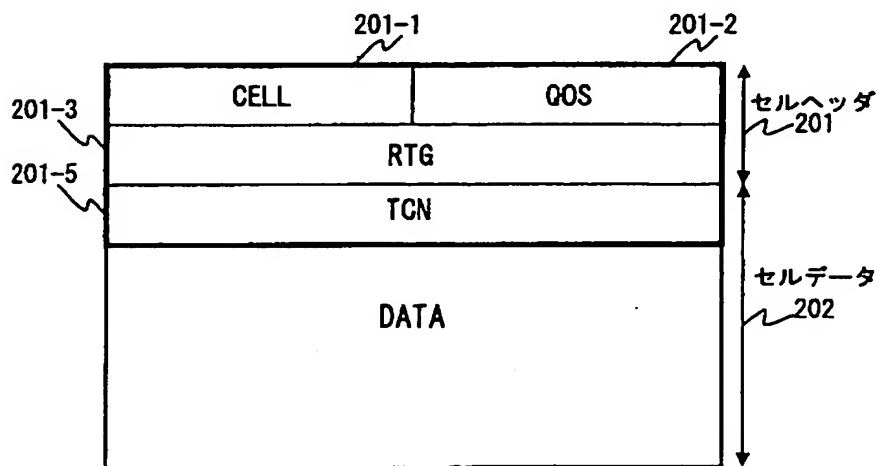
【図 13】

図 13



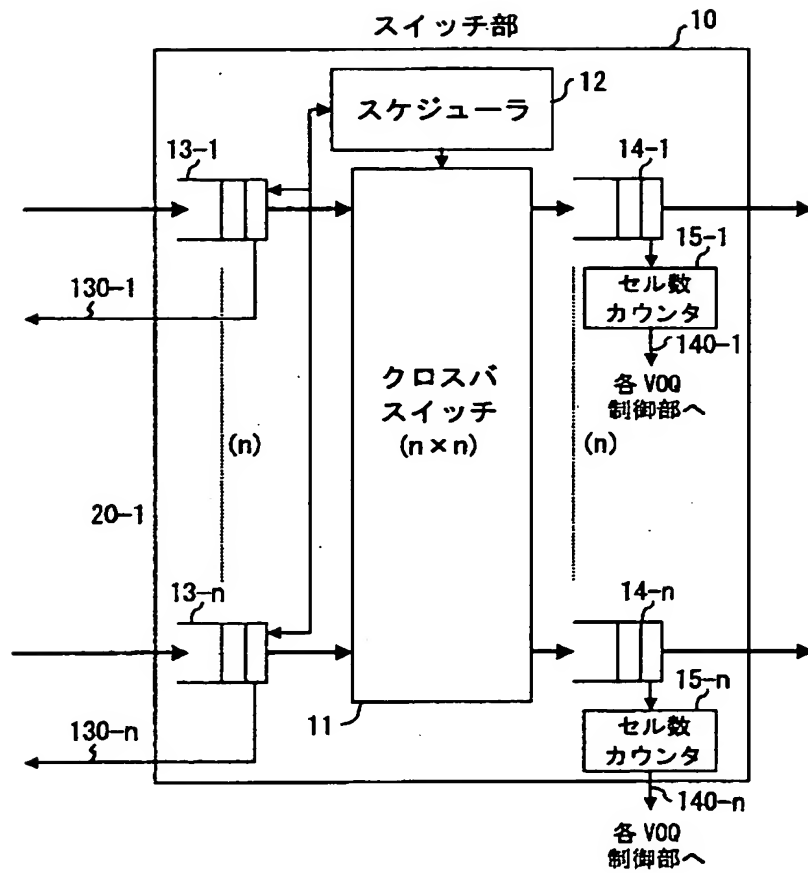
【図 14】

図 14



【図 15】

図 15



【書類名】 要約書

【要約】

【課題】 可変長パケット単位でスイッチの入出力ポートの接続を変更するパケットスイッチにおいて、高スループットおよびパケットの優先制御が実現可能な大容量のパケット通信装置を提供することである。

【解決手段】 入力インタフェースにおいて可変長パケットをセル群に分割し、スイッチの宛先出力ポート別に分割されたVOQに格納する。各VOQには、それぞれに対応した先頭セル格納レジスタが設けられ、パケットがVOQの先頭に到着するとそのパケットの出力方路を表示した先頭セルを、先頭セル格納レジスタに転送する。各入力インタフェースは出力可能なパケットの先頭セルを1つ選択してスイッチへ送信する。スイッチでは、出力ポートにつき1つの先頭セルを選択するようにスケジューリング処理を行う。スケジューリング処理の結果、出力許可された入力インタフェースは、所望の出力ポートに接続され、先頭セルおよびVOQに格納されている後続セルをパケット単位で出力ポートに連続的に出力する。

【選択図】 図4

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日
[変更理由] 新規登録
住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所